

Project plan

Dhanush Kikkiseti & Druv Jain

2022-10-18

Plan 1. Analysis Title and Type of Analysis

Analysis Title - Analysing **Global Power Plant** dataset **Type of Analysis** - Research Project(Observational Study)

Plan 2 Names and Courses of your group members

Dhanush Kikkiseti - DATA-612
Druv Jain - Data-612

Plan 3 descriptions of problem to be analyzed

This dataset consists of electricity generated by different countries in the globe using different powerplant (Hydro, Gas, Oil, solar, wind, nuclear, biomass, waste) from year 2013 to 2018.



There are many problems that can be taken into consideration. We can frame more problems when we start working on data.

Let's look at some problems that can be analyzed

- 1) Choosing a particular country and analyzing what type of powerplant has more electricity generation .
- 2) Analyzing which type of powerplant has more electricity generation and doing a little research on it.
- 3) Analyzing if there is a difference in generated electricity and estimated electricity with respect to the country.
- 4) Effect of power plant type in generating electricity.
- 5) Comparing electricity generation in different years with respect to country

Plan 4. Proposed data sources

[1] website name: **WORLD RESOURCES INSTITUTE** 

[2] Website link: Dataset page

[3] Citation: Global Energy Observatory, Google, KTH Royal Institute of Technology in Stockholm, Enipada, World Resources Institute. 2018. Global Power Plant Database. Published on Resource Watch

Plan5 Proposed approach for the Analysis

- 1) There are a few things to be considered before solving the data. One of them is that it is data structured or unstructured. Secondly, making various assumptions and removing some non-obligatory data. Gathering and performing some data cleansing operations to make look data clean and useful for visualization and performing statistical computations for our question of interest.
- 2) After data is cleaned, using particular attributes in the dataset for visualizations like
Making a box plot for analysis of electricity generation for 5 different years.
Histogram for checking how generated electricity data is distributed with respective country
Scatterplot for estimated and generated electricity
ETC....
- 3) Using statistical methods for checking a hypothesis (question of interest) for the dataset

Plan 6. Allocation of Responsibilities for the team

Dhanush will take care of data cleansing, EDA and other stats analysis works

Dhruv jain will take care report making and performing EDA

Plan 7. Method and Timing for Collaboration

Method: In-person or in case of not availability will take this to zoom call.

Timing for collaboration: We are planning to give 4 hours of work every week to make sure that all the responsibility has been taken care of with proper time management. Majorly we are planning to work Wednesday from 4:00 pm till 6 pm and on Fridays we are planning to work between 1 to 3 pm at noon time.

Plan 8. Project Steps/Schedule

No.	Start date	Close Date	TO DO BEFORE THAT DAY
1	10/13/2022	10/18/2022	Submit the rough project plan
2	10/19/2022	10/27/2022	Gathering of Data, cleaning
3	10/28/2022	11/02/2022	New data analyzing, problem identifying
4	11/03/2022	11/10/2022	Working on problems
5	11/11/2022	11/17/2022	Statistical analysis
6	11/18/2022	11/24/2022	Comparing prediction cs generated values
7	11/25/2022	12/01/2022	Graphical visualizations, report making
8	12/02/2022	12/08/2022	Final touch before submitting, ppt making
9	12/09/2022	12/15/2022	Final report & presentation